



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

MOOCs in the news

Citation for published version:

Kovanovic, V, Joksimovic, S, Gasevic, D, Hatala, M & Siemens, G 2015, MOOCs in the news: A European perspective. in *Proceedings of the WOW! Europe embraces MOOCs Conference*.

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Published In:

Proceedings of the WOW! Europe embraces MOOCs Conference

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



MOOCs in the News: A European Perspective

Vitomir Kovanović¹, Dragan Gašević^{1,2}, Srećko Joksimović², George Siemens³, [Marek Hatala](mailto:Marek.Hatala@ed.ac.uk)⁴

¹School of Informatics, University of Edinburgh
(v.kovanovic@ed.ac.uk)

²Moray House School of Education, University of Edinburgh

³Link Research Lab, University of Texas at Austin

⁴School of Interactive Arts and Technology, Simon Fraser University

Abstract

Recent development of Massive Open Online Courses (MOOCs) commenced unprescedented interest of the general public. To leverage from attention given to MOOCs, understanding of public discourse is essential, as it can give critical insights into the important domains of biggest societal interests. Prevoius research showed the great need for understanding specifics of MOOC adoption around the world and the necessity to better cater to the needs of different markets. With this in mind, this paper presents study that looked specificially at the Europe-related MOOC discourse between 2008 and 2015. We identified important themes in the MOOC public discourse and evaluated their changes over time. Further implications of our findings are also discussed.

Keywords

MOOCs, MOOC public discourse, topic modeling, MOOCs in Europe

1 Introduction

Although there have been many advances in the educational technology field over the years, the recent development of Massive Open Online Courses (MOOCs) is particularly interesting given its large coverage in the mainstream media (Stewart, 2013a). From a small initiative by a small group of educational technology researchers, MOOCs become one of the most prominent educational technology topics (Siemens, 2012), often described as disruption or revolution in education (Hennessy, 2012). This high interest by the general public is likely due to the combination of different social, political, and economical reasons (Bates, 2014) including involvement of elite universities and Silicon Valley companies, and the overall economic climate after the 2008 financial crisis.

Although MOOCs were significantly covered by the mainstream media since the announcement of the first Stanford MOOCs, the tone and focus of those media reports changed substantially. Initially, MOOCs were presented as revolution in education (Friedman, 2012; Hennessy, 2012), with the famous New York Times article labeling 2012 as the "year of the MOOCs" (Pappano, 2012). Later on, the discussion became more critical, with some suggesting that MOOCs failed to reach their promises (Adams, 2013; Stober, 2015), whereas others suggest that passing of the "MOOC hype" brings more productive conversation regarding the position of MOOCs in the broader landscape of education (Lewin, 2013; Muldowney, 2015; Oxenham, 2015; Stewart, 2013b). Given the large adoption of online and distance education, it is important to understand these changes in the MOOC media image in order to capitalize on the present momentum that MOOCs have brought to the educational technology domain.

The goal of this paper is to investigate the public discourse surrounding MOOCs since the first offerings of the MOOCs to the present day. Building on our previous study (Kovanović, Joksimović, Gašević, Siemens, & Hatala, 2015) that investigated global trends in MOOC public discourse, the focus of this paper is particularly on the European perspective of MOOC media coverage. Through a systematic search

of news reports, we identified 915 news reports between January 2008 and November 2015 related to MOOCs in Europe. We conducted an automated topic modeling analysis using Latent Dirichlet Allocation (LDA) technique which produced 43 distinctive topics. In this paper we focus on top twenty most prominent topics and also investigate their change over time. The interpretation of the findings and their implications are further discussed.

2 Background

2.1 Analysis of MOOC research literature

Although MOOCs are still in early days of their development, there have been several attempts to examine current literature related to MOOC research and practice. Analysis by Liyanagunawardena, Adams, and Williams (2013) looked at the published MOOC research literature, and revealed a strong focus on students' learning experience and institutional adoption opportunities and challenges. Despite various learning and social media platforms being used by the learners, researchers typically focus only on a small portion of the data coming from a single platform – as collecting and linking data from different software platforms is often very challenging – which limits the understanding of the MOOC learning processes (Liyanagunawardena et al., 2013). Furthermore, a study by Gašević, Kovanović, Joksimović, and Siemens (2014) looked at the state of the MOOC research as reflected by the submissions to the MOOC Research Initiative (MRI) (MOOC Research Initiative, 2013), revealing a fragmentation in the MOOC research community – one group of mostly educational researchers gathered around MRI initiative, and the second group of predominantly computer-science researchers gathered around ACM Learning@Scale conference. Gašević et al. (2014) results also revealed five important areas of MOOC research: i) student engagement and learning success, ii) MOOC design and curriculum, iii) self-regulated learning and social learning, iv) social network analysis and networked learning, and v) motivation, attitude and success criteria.

2.2 Analysis of MOOC public media

Besides investigation of MOOC research literature, there have been several studies looking explicitly at the public media discourse relating to MOOCs. The analysis by Selwyn, Bulfin, and Pangrazio (2015) and Bulfin, Pangrazio, and Selwyn (2014) of 457 MOOC-related news reports reveal the focus on marketization, monetization, and massification aspects of MOOCs, rather than the debate on the pedagogical or technological aspects of MOOC course design or student learning experience. The primary themes in MOOC public discourse are related to i) MOOCs bringing change to education landscape, ii) MOOCs being free of charge, and iii) MOOCs being large scale (Bulfin et al., 2014).

In their analysis of 4024 MOOC-related news articles from around the world, Kovanović et al. (2015) identified important themes in MOOC public discourse and their change over time. Kovanović et al. (2015) study revealed rapid decrease in MOOC news coverage and the move from provider-focused discussions to more productive discussions centered around position of MOOCs in the global educational landscape, the use of big data and analytics, and government-related regulations. The focus of current criticism of MOOCs is primarily on the failure of MOOCs to bring "the revolution" to the field of education (Kovanović et al., 2015). Finally, there has been a growing number of topics related to MOOC adoption around the world, showing the need to better cater MOOCs to the needs of different markets. With this in mind, the goal of this paper is to examine MOOC media coverage in Europe and identify prominent themes in the discourse, and their changes over time.

3 Method

3.1 Dataset

Similarly to our previous work (Kovanović et al., 2015), the data for this study is obtained through Factiva (Dow Jones & Company, 2014), which is a business information retrieval tool developed by Dow Jones & Company and Reuters news

agency. Factiva is one of the largest databases of news articles, containing millions of both free and licensed news articles from around the world (Dow Jones & Company, 2014). We conducted a search for news articles written in English and containing “MOOC(s)” or “Massive Open Online Course(s)” keywords. To remove irrelevant results, we limited our search to the Europe-related news articles published between Jan 1, 2008 and Nov 15, 2015. In total, we obtained 974 search results which were then downloaded and further examined. As Factiva also contains different types of documents besides news articles, we manually examined the search results and removed the irrelevant documents which resulted in 915 articles being finally included in our dataset.

3.1 Analysis procedure

To analyze our dataset, we used Latent Dirichlet Allocation (LDA) (Blei, Ng, & Jordan, 2003), a popular probabilistic topic modeling technique used to identify prominent themes in the document corpora. We implemented our analysis in R programming language (R Core Team, 2013) and `topicmodels` LDA library (Grun & Hornik, 2014). LDA works by looking at the co-occurrence of words in the dataset, finding groups of words that are frequently used together and represent a distinct topic in the corpora. It is often used for analysis of large bodies of text in social sciences and humanities (Cohen et al., 2012), including the analyses of news articles (Wei & Croft, 2006; Yang, Torget, & Mihalcea, 2011).

The main input to LDA is the document-term matrix (DTM) which is a matrix indicating how many times each unique word appears in all documents in the corpora. Before running LDA algorithm, we preprocessed the data by i) removing stop-words (i.e., very frequently occurring words such as ‘a’, ‘the’, ‘be’, ‘of’ etc.), ii) removing numbers and URLs, iii) removing short words (i.e., less than 3 characters long), and iv) word lemmatization (i.e., reducing words to their base forms, for example ‘walking’ to ‘walk’). This resulted in 15,882 unique terms being extracted. However, given that most words only appear in a tiny fraction of documents, we removed all terms (i.e., columns) that appear in less than 5% of the

documents (i.e., rows). This resulted in reducing the number of extracted terms down to 1,114 terms which improves the quality of the extracted topics as the data scarcity negatively impacts topic extraction procedure ([Hong & Davison, 2010](#)). Finally, after removing very rare terms, we also removed frequent, but non-important words that are not useful for topic extraction procedure. In the similar manner as in the Kovanović et al. (2015) study, we removed words with their TF-IDF score below 0.95 of the median TF-IDF value. This further reduced down our number of terms to 732 terms.

Besides document-term matrix, LDA requires the number of topics to be defined in advance. Since we do not know how many topics are in the data, we evaluated all LDA models with 2 to 100 topics, and used maximum likelihood method described by Ponweiser (2012) to select the optimal number of topics. Given that this requires evaluation of the large number of LDA models, we used randomly selected 20% of the data as input to LDA procedure. After the optimal number of topics is extracted, the new model on the full data is fitted and analyzed.

4 Results

4.1 Data collection results

Table 1 shows the number of articles across the covered years (2008-2015). We can see that before 2012, there were only 2 news articles related to MOOCs in 2009. Thus, in the reminder of this paper, we focus our investigation only on articles published between 2012 and 2015. With respect to article length, Figure 1 shows the number of words per article. The distribution of article lengths reasonably follows the normal distribution, with only a longer tail on the right (as article length cannot be negative). The average number of words per article is 716 words, and that majority of articles have between 400 and 900 words. This is very similar to the average article length of 765 reported by Kovanović et al. (2015) and not much different from the average news article length of 800 words, as reported by Project for Excellence in Journalism (2004).

Table 1: Numbers of articles per year for the period covered by the study

Year	Article count	Percentage of total
2008	0	0.00
2009	2	0.00
2010	0	0.00
2011	0	0.00
2012	76	0.08
2013	341	0.37
2014	296	0.32
2015 (up to Nov 15)	200	0.22
Total	915	1.00

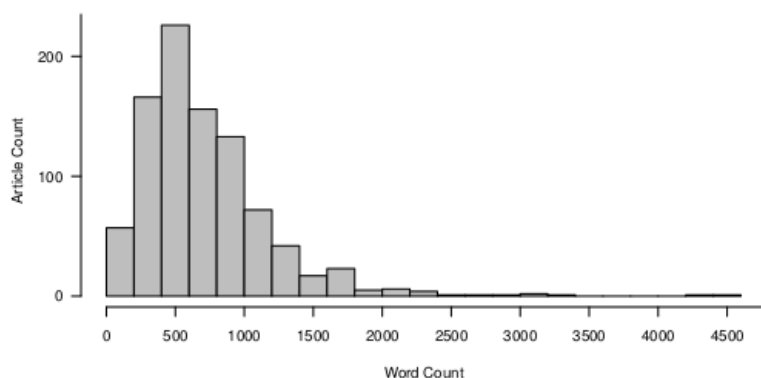


Figure 1: Number of words across all news articles. Mean(SD)=716(475) words;
Median(Q1,Q3)=613(406,914) words.

Looking at sources of Europe-related MOOC news articles (Table 2), we see that news publishers from UK were most dominant, which is not surprising given our focus on news articles written in English. Aligned with Kovanović et al. (2015) results, we find that most news source published between one and three MOOC-related articles. The most articles (99) were published by Times Higher Education Supplement, which is twice as much as the number of articles (49) published by Financial Times, the second most frequent news source. As expected, the biggest interest in MOOCs is shown by education-related publications, and large national newspapers from the UK. Interestingly, several US and Australian news agencies covered European MOOC-related news. We also see interest in MOOCs by news publishers focused primarily on business and finance (i.e., Financial Times, Mena Report, Australian Financial Review) which is the trend also witnessed in the previous studies (Kovanović et al., 2015).

4.2 Topic modeling results

In order to select the optimal number of topics, we evaluated all topic modeling solutions having between two and hundred topics (Figure 3). Our analysis identified solution with 43 topics as the optimal one, which was the one that was used in the reminder of this paper. We used 43 topics to fit the model on the all 915 articles and assigned each article to one of the 43 identified topics, based on the assigned log-likelihoods (Figure 4). In rare cases where it was equally likely that a given article belongs to two or more topics, we assigned a given article to all of the most likely topics found. Looking at the Figure 4, we can see the steep decline after the top three topics – indicating that they were significantly more covered in the news than other topics – while the remaining topics show much more monotonic decrease in the coverage. With the goal of covering as much relevant themes as possible in the limited space, in the reminder of this paper we focused on the top twenty most prominent topics. Overall, top twenty topics cover just slightly below two thirds of the 915 articles in our dataset. To describe each of the topics, we looked at the assigned documents and the list of words mostly associated with each of the topics.

Table 2: Twenty most prominent news sources

#	Source	Articles	#	Source	Articles
1	The Times Higher Ed. Supp.	99	11	The Irish Times	18
2	Financial Times (FT.Com)	49	12	The Sunday Times	18
3	The Guardian	47	13	PR Newswire (U.S.)	16
4	Financial Times	31	14	PR Newswire Europe	14
5	Press Assoc. National Newswire	30	15	Mena Report	13
6	The Telegraph Online	30	16	The Independent	13
7	M2 Presswire	23	17	The Times	12
8	ENP Newswire	22	18	Independent Online	11
9	Education Letter	21	19	Australian Financial Review	11
10	Guardian.co.uk	18	20	Sundaytimes.co.uk	10

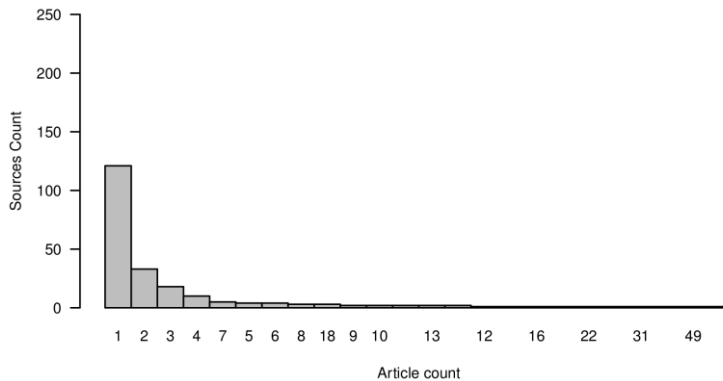


Figure 2: Number of articles across all news sources included in the study.
Mean(SD)=9(9) articles; Median(Q1, Q3)=1(1,3) articles.

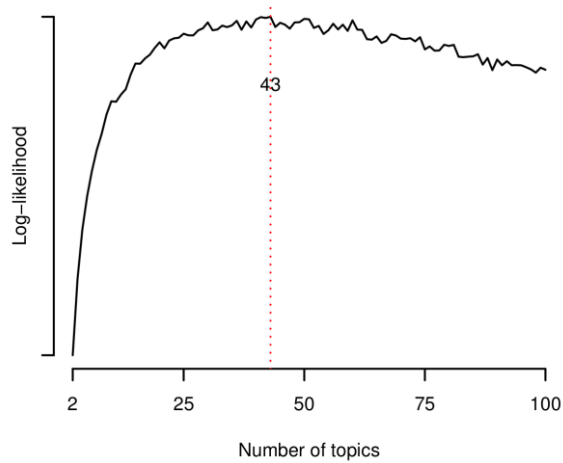


Figure 3: Log-likelihood of different topic solutions.

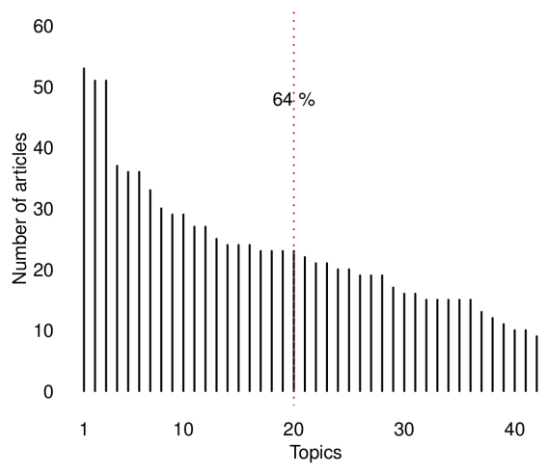


Figure 4: Number of documents per each topic.

Table 3 shows the list of top ten most relevant terms for the twenty most prominent topics. Based on the identified terms and associated news articles, we defined labels for each of the twenty topics. Looking at the Table 3, we can see that three topics that were mostly covered in published news sources are related to i) FutureLearn, a UK based MOOC platform supported by the Open University UK, ii) business and management (MBA) MOOC offerings, and iii) the “MOOC revolution” led by the Coursera, Udacity, EdX and elite institutions such as Stanford University. Besides these three topics, commonly discussed topics are related to:

- Use of MOOCs for K-12 education (i.e., primary and secondary education),
- Changes in university funding due to wide availability of MOOC courses,
- MOOCs in Ireland,
- Announcements of different MOOC courses,
- The impact of MOOCs on the global educational market
- Press releases related to openSAP, the MOOC platform developed by SAP,
- MOOC signup figures, primarily related to UK universities,
- Introductions to MOOC courses and mechanics of online education,
- Articles discussing different research reports related to MOOC space,
- MOOC in France,
- Development of books and written materials for MOOC contexts,
- Use of MOOC for workplace training,
- Changes in the private educational sector caused by MOOCs,
- Announcements of British Council MOOC for teaching English language,
- MOOC market,
- Different European initiatives related to MOOCs, and
- Announcements of MOOC-related conferences.

Table 3: Ten most relevant terms for the twenty most prominent topics

#	Topic Label	Articles	Distinctive Terms
1	FutureLearn	53	futurelearn, nelson, partner, leed, warwick, bbc, learner, simon, east, bristol
2	Business and Management MOOCs	51	business, school, mba, management, dean, finance, malster, prof, manager, case
3	MOOC revolution	51	coursera, stanford, credit, udacity, sign, certificate, completion, venture, enrol, edx
4	MOOCs for K-12 education	37	school, teacher, pupil, lesson, child, computing, resource, curriculum, classroom, secondary
5	MOOCs and university funding	36	government, funding, tuition, sector, pound, loan, overseas, fall, private, target
6	MOOCs in Ireland	36	distance, irish, ireland, trinity, postgraduate, dublin, tutor, history, qualification, law
7	MOOC course announcements	33	vice, chancellor, david, bean, minister, willett, martin, sir, december, widen
8	MOOCs and global educational market	30	china, profit, american, campus, elite, america, accord, emerge, expand, mass
9	openSAP press releases	29	com, www, http, press, visit, newswire, release, solution, announce, business
10	MOOC signup figures for UK universities	29	london, edinburgh, join, sign, oxford, king, principal, australia, september, vice
11	Explaining MOOC course structure	27	video, forum, website, peer, assignment, exam, grade, youtube, user, quiz
12	MOOCs research reports	27	news, report, accord, additional, contact, editor, obtain, article, journal, learner
13	MOOCs in France	25	edx, platform, france, french, national, source, october, january, announce, massachusetts

14 MOOCs & books	24	book, story, claim, tell, hour, read, word, history, predict, educate
15 MOOCs for training	24	skill, training, employer, career, qualification, professional, scheme, workplace, employee, national
16 MOOCs & changes in educational sector	24	per, cent, service, face, sector, survey, process, average, march, office
17 British council MOOCs	23	british, council, library, museum, futurelearn, debate, english, announce, activity, cultural
18 MOOC market	23	market, industry, report, trend, investment, sector, analysis, growth, company, product
19 European MOOC initiatives	23	europe, european, resource, commission, initiative, innovative, unit, ensure, formal, adopt
20 MOOC conferences	23	conference, event, educator, foundation, contact, present, communication, worldwide, meeting, discuss

Besides identifying most frequently discussed topics, we examined the dynamics of their coverage over time. Figure 5 shows the changes in coverage for the top twenty topic across the 2012-2015 period. Interestingly we see the steep decline for the large number of topics, in particular the top three most discussed topics. Although there is an overall decline for the large number of topics, some – such as the coverage of MOOC reports, university funding, openSAP, use of books, use of MOOCs for training, and MOOC market show increase over time.

5 Limitations

There are several limitations related to our study. First of all, we used the data that was available on the Factiva information retrieval platform, which – despite being one of the largest databases available – is still not complete set of MOOC-related articles. This is particularly true for the articles published in 2015, as the Factiva database is slightly lagging behind the actual news publishers. Next, al-

though the use of fully automated analysis procedure has many advantages including the ability to analyze hundreds of news articles in a quick and easy manner, the level of sophistication that can be achieved by the expert researcher is still higher. Thus, similarly to Kovanović et al. (2015) study, we argue that combination of smaller but more detailed analysis – such as the ones by Bulfin et al. (2014) and [Selwyn et al. \(2015\)](#) – and the automated analysis like the one presented here provide a necessary insights to the complex nature of MOOC public discourse. The adopted procedure also depends on the several preprocessing steps and algorithm parameters, the primary one being number of topics to extract which both can have an important impact on the final results of the analysis

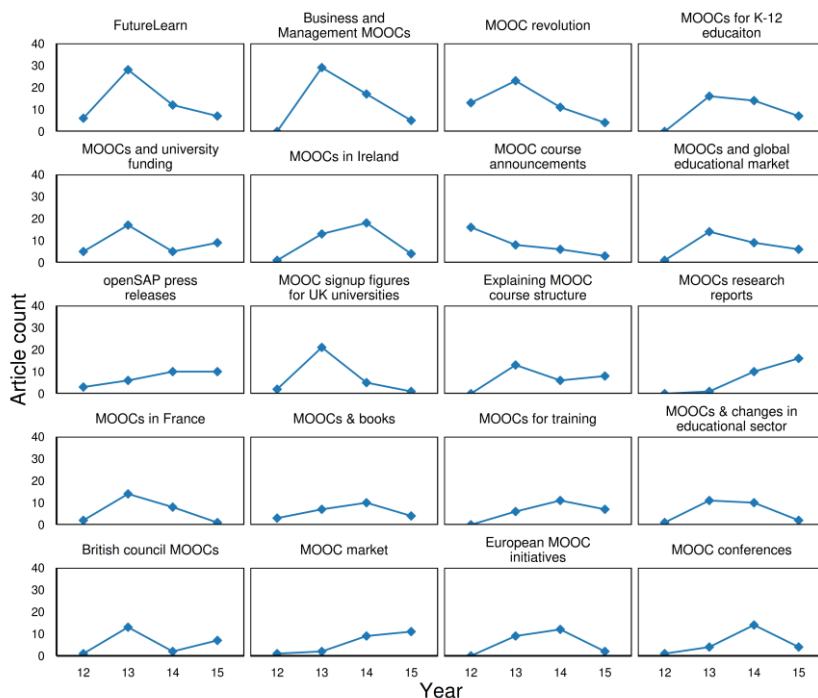


Figure 5: Change in coverage of top twenty topics over 2012-2015 period.

Finally, although having focus on the whole Europe, we used only articles written in English, which distorts a final image to more UK- and Ireland-related topics. In our future work, we plan on investigating articles written in other European languages to provide a more comprehensive overview of the European MOOC public discourse. One promising direction is to use automated translation services such as Google Translate to translate all articles to English before running topic modeling procedure. Although not perfect, those automated translation services could provide translation which is accurate enough so that an accurate document-term matrix could be extracted.

6 Discussion

Looking at the results of our analysis, we see many similarities with the previous results reported by Kovanović et al. (2015). As expected, a large portion of public discourse has been about “MOOC revolution” and the anticipated changes in the educational domain. Also, given the large amounts of money being raised by MOOC companies, we also witness many financial and business-oriented publications reporting extensively on MOOCs affairs. This also explains why MOOC providers and business and management MOOCs are extensively covered in the public media. This focus on financial aspects of MOOC triggers a question of how much research in online learning is underpinning the development in MOOC space, and how much it is driven by the marketability, particularly in the case of for-profit MOOC providers (Kovanović et al., 2015).

We also see the decreasing trend in MOOC coverage over time, albeit to the lesser extent than reported by Kovanović et al. (2015). One likely reason is slightly slower adoption of MOOCs in Europe which resulted in peek of “MOOC hype” cycle being shifted. Thus, based on results of this study and Kovanović et al. (2015) study to continue to see a declining trend in MOOC coverage in Europe in the following years. We also see a number of topics related to MOOC use in UK, Ireland, and France, which is aligned with the results by [Selwyn et al. \(2015\)](#) and Kovanović et al. (2015). We also see large coverage of smaller MOOC platforms,

such as openSAP from non-English speaking countries, which is also an indicator of the importance of MOOC adoption to the different user populations, markets, and economies.

Although large number of topics saw decline in their coverage, certain number of topics show trend of rising coverage over time. For instance, changes in university funding due to the introduction of MOOCs saw a decline between 2013 and 2014, but saw a second increase during 2015. As MOOCs became more mature, there has also been an increase in the number of MOOC-related reports, development of different MOOC books and learning materials, and discussions related to the use of MOOCs in the context of workplace training. This indicates that the public discourse of MOOCs is moving towards more productive analysis of instructional aspects of MOOCs and the positions that MOOCs might play in the overall educational landscape and the need of current universities to adjust to this new market organization.

7 Conclusions

This paper presents a study that looked at the Europe-related MOOC public discourse. We analyzed 915 English language news articles obtained through Factiva platform using automated topic modeling technique previously used by Kovanović et al. (2015). Unsurprisingly, the most MOOC-related articles (99) have been published by Times Higher Education Supplement. Besides education-focused publications, several financial and business newspapers extensively reported on European MOOC affairs. Alongside several UK daily and weekly newspapers, we see large coverage of European MOOC news in US and Australian newspapers. Using Latent Dirichlet Allocation (LDA), we found 43 distinct topics in our dataset. Aligned with Kovanović et al. (2015) results, we see a move from broad discussion of MOOCs and MOOC providers to more constructive discussion related to MOOC adoption and their position in the educational field. The three most prominent topics were related to FutureLearn MOOC platform, Business and Management MOOCs, and “MOOC revolution”, which all suffered large drop in the coverage in

the last two years. In contrast, we see increase in coverage of some other topics, such as the discussion of MOOC research reports, analysis of MOOC markets, use of MOOCs in primary and secondary education, and changes to university funding initiated by MOOCs.

References

- Adams, S. (2013). Are MOOCs Really A Failure? *Forbes*. Retrieved from <http://www.forbes.com/sites/susanadams/2013/12/11/are-moocs-really-a-failure/>
- Bates, T. (2014, November 16). Why the fuss about MOOCs? Political, social and economic drivers. Retrieved from <http://www.tonybates.ca/2014/11/16/why-the-fuss-about-moocs-political-social-and-economic-drivers/>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *J. Mach. Learn. Res.*, 3, 993–1022.
- Bulfin, S., Pangrazio, L., & Selwyn, N. (2014). Making “MOOCs”: The construction of a new digital higher education within news media discourse. *The International Review of Research in Open and Distance Learning*, 15(5). Retrieved from <http://www.irrodl.org/index.php/irrodl/article/view/1856>
- Cohen, D. J., Troyano, J. F., Hoffman, S., Wieringa, J., Meeks, E., & Weingart, S. (Eds.). (2012). Special Issue on Topic Modeling in Digital Humanities. *Journal of Digital Humanities*, 2(1).
- Dow Jones & Company. (2014). Factiva business information tool. Retrieved from www.factiva.com

Friedman, T. L. (2012, May 15). Come the Revolution. *The New York Times*. Retrieved from <http://www.nytimes.com/2012/05/16/opinion/friedman-come-the-revolution.html>

Gašević, D., Kovanović, V., Joksimović, S., & Siemens, G. (2014). Where is research on massive open online courses headed? A data analysis of the MOOC Research Initiative. *The International Review of Research in Open and Distributed Learning*, 15(5). Retrieved from <http://www.irrodl.org/index.php/irrodl/article/view/1954>

Grun, B., & Hornik, K. (2014). [topicmodels: An R Package for Fitting Topic Models](#).

Hennessy, J. L. (2012). *The Coming Tsunami in Educational Technology*. Presented at the The 40th CRA Conference at Snowbird.

Hong, L., & Davison, B. D. (2010). Empirical Study of Topic Modeling in Twitter. In [Proceedings of the First Workshop on Social Media Analytics](#) (pp. 80–88). New York, NY, USA: ACM. <http://doi.org/10.1145/1964858.1964870>

Kovanović, V., Joksimović, S., Gašević, D., Siemens, G., & Hatala, M. (2015). What public media reveals about MOOCs: A systematic analysis of news reports. *British Journal of Educational Technology*, 46(3), 510–527.

Lewin, T. (2013, December 10). After Setbacks, Online Courses Are Rethought. *The New York Times*. Retrieved from <http://www.nytimes.com/2013/12/11/us/after-setbacks-online-courses-are-rethought.html>

Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: A systematic study of the published literature 2008-2012. *The International Review of Research in Open and Distance Learning*, 14(3), 202–227.

MOOC Research Initiative. (2013). www.moocresearch.com.

Muldowney, S. (2015). Out-moocing the MOOCs: Traditional institutions fight back. Retrieved from <http://intheblack.com/articles/2015/11/01/out-moocing-the-moocs-traditional-institutions-fight-back>

Oxenham, S. (2015). Can Online Learning Ever Beat the Real Thing? *Big Think*. Retrieved from <http://bigthink.com/neurobonkers/can-online-learning-ever-beat-the-real-thing>

Pappano, L. (2012, November 2). The Year of the MOOC. *New York Times*. Retrieved from <http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html?pagewanted=all>

Ponweiser, M. (2012). *Latent Dirichlet Allocation in R* (Diploma Thesis). Vienna University of Economics and Business, Vienna.

Project for Excellence in Journalism. (2004). *The State of the News Media*. The Pew Research Center's Journalism Project.

R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Selwyn, N., Bulfin, S., & Pangrazio, L. (2015). Massive Open Online Change? Exploring the Discursive Construction of the "MOOC" in Newspapers. *Higher Education Quarterly*, 69(2), 175–192. <http://doi.org/10.1111/hequ.12061>

Siemens, G. (2012). MOOCs are really a platform. *ELearnSpace Blog*. Retrieved from <http://www.elearnspace.org/blog/2012/07/25/moocs-are-really-a-platform/>

Stewart, B. (2013a). Massiveness + Openness = New Literacies of Participation? *Journal of Online Learning and Teaching*, 9(2), 44–56.

Stewart, B. (2013b). the post-MOOC-hype landscape: what's REALLY next? Retrieved from <http://theory.cribchronicles.com/2013/12/10/the-post-mooc-hype-landscape-whats-really-next/>

Stober, D. (2015, October 15). MOOCs haven't lived up to the hopes and the hype, Stanford participants say. *Stanford University*. Retrieved from <http://news.stanford.edu/news/2015/october/moocs-no-panacea-101515.html>

Wei, X., & Croft, W. B. (2006). LDA-based Document Models for Ad-hoc Retrieval. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 178–185). New York, NY, USA.

Yang, T.-l., Torget, A. J., & Mihalcea, R. (2011). Topic Modeling on Historical Newspapers. In *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities* (pp. 96–104). Stroudsburg, PA, USA.



This work is licensed under a Creative Commons Attribution 4.0 International License:

<https://creativecommons.org/licenses/by/4.0/> This li-

cense lets others distribute, remix, tweak, and build upon this work (even commercially) as long as credit is provided for the original creation. This is among the most accommodating of CC licenses offered, and recommended for maximum dissemination and use of licensed materials.